

Deep learning based pulse prediction of nonlinear dynamics in fiber optics

*Original*

Deep learning based pulse prediction of nonlinear dynamics in fiber optics / Sui, H.; Zhu, H.; Cheng, L.; Luo, B.; Taccheo, S.; Zou, X.; Yan, L.. - In: OPTICS EXPRESS. - ISSN 1094-4087. - ELETTRONICO. - 29:26(2021), pp. 44080-44092. [10.1364/OE.443279]

*Availability:*

This version is available at: 11583/2959935 since: 2022-03-29T16:16:00Z

*Publisher:*

Optica Publishing Group

*Published*

DOI:10.1364/OE.443279

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

Optica Publishing Group (formerly OSA) postprint versione editoriale con OAPA (OA Publishing Agreement)

© 2021 Optica Publishing Group. Users may use, reuse, and build upon the article, or use the article for text or data mining, so long as such uses are for non-commercial purposes and appropriate attribution is maintained. All other rights are reserved.

(Article begins on next page)



# Deep learning based pulse prediction of nonlinear dynamics in fiber optics

HAO SUI,<sup>1</sup>  HONGNA ZHU,<sup>1,\*</sup>  LE CHENG,<sup>1</sup> BIN LUO,<sup>2</sup> STEFANO TACCHEO,<sup>3</sup> XIHUA ZOU,<sup>2</sup>  AND LIANSHAN YAN<sup>2</sup> 

<sup>1</sup>*School of Physical Science and Technology, Southwest Jiaotong University, Chengdu 610031, China*

<sup>2</sup>*School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China*

<sup>3</sup>*Department of Electronics and Telecommunications, Politecnico di Torino, Torino 10129, Italy*

\*[honzhu@swjtu.edu.cn](mailto:honzhu@swjtu.edu.cn)

**Abstract:** The initial state of a nonlinear optical fiber system plays a vital role in the ultrafast pulse evolution dynamic. In this work, a data-driven compressed convolutional neural network, named inverse network, is proposed to predict initial pulse distribution through a series of discrete power profiles at different propagation distances. The inverse network is trained and tested based on two typical nonlinear dynamics: (1) the pulse evolution in a fiber optical parametric amplifier system and (2) soliton pair evolution in high-nonlinear fibers. Great prediction accuracy is reached when the epoch grows to 5000 in both cases, with the normalized root mean square errors below 0.01 on the entire testing set. Meanwhile, the lightweight network is highly effective. In this work, it takes approximately 30 seconds for 5,000 epochs training with a dataset size of 900. The inverse network is further tested and analyzed on the dataset with different signal-to-noise ratios and input sizes. The results show fair stability at the deviation on the testing set. The proposed inverse network demonstrates a promising approach to optimizing the initial pulse of fiber optics systems.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

The propagation of ultrashort pulse in optical fiber, which exhibits highly nonlinear evolution, has attracted attentive research interests and wide applications in the field of nonlinear fiber optics [1–8]. With the continuous development of optical fiber system towards ultra-high speed, ultra-large capacity and ultra-long distance, the utilization and suppression of nonlinear pulse propagation in optical fiber is playing the governing roles in a nonlinear fiber system [9]. However, the complex evolutions make it challenging and time-consuming to parameters optimization and control these dynamics by conventional numerical methods [10].

As a powerful tool for system parameter optimization and construction of models of complex dynamics from observed data, deep learning (DL) algorithms have been recently applied to ultrafast photonics [11], optical networks [12] as well as other optical fiber systems to add new functionalities and enhance performance [13–17]. There is recent literature of solving the nonlinear pulse propagation governed by the nonlinear Schrödinger equation (NLSE) via DL methods.

For example, the recurrent neural network has been applied in predicting complex nonlinear propagation in both high-nonlinear fiber (HNLF) and multimode fibers [10,18], as well as modeling optical fiber channel [19]. Physics-informed neural network was proposed to solve the NLSE [20]. Deep learning approach was also used to simulate rogue waves [21] and self-similar parabolic pulses [22–23] governed by NLSE. The comparison of different neural network architectures for modelling nonlinear pulse propagation was also investigated [24]. The mentioned works focused on the forward propagation problem of the modeling of the complex nonlinear optical dynamics from a given input pulse. The corresponding inverse propagation problem that assesses the initial pulse parameters or distributions from its nonlinear evolution

is also of vital importance and has potential applications in the design and optimization of experiments and real-time optical systems.

The digital backpropagation (DBP) algorithm is a promising method to the inverse propagation problem which works by backpropagating the received signal with inverted fiber parameters. However, this method requires high computational resources and full knowledge of system parameters. In addition, the accuracy of the results cannot be guaranteed due to the presence of random parameters [25–26]. In the system where a slight change in one parameter may lead to great deviation to pulse evolution, the data-driven methods based on DL can be superior and desirable. Therein the networks can learn from the received data and counteract the effects of nonlinearity without prior knowledge. System agnostic nonlinear impairment compensation was achieved by neural network [26]. The convolutional neural networks were used for compensating nonlinear distortions in a long-haul ultrahigh capacity fiber-optic transmission system [27]. Different neural networks for obtaining the output pulse characteristics and reconstructing the input pulse were achieved and compared [24]. The soliton properties were estimated from the nonlinear propagation map by two types of neural networks [28].

Note that the inverse propagation problems in other nonlinear dynamics are extremely desirable for expending the applications of DL methods in optimizing of NLSE-based systems and overcoming time-consuming numerical simulations in a conventional optimization. Also, it is necessary to investigate the network on the datasets different from the training set (i.e., dataset with different fiber parameters, different signal-to-noise ratios and input size) to show the stability of the network and to demonstrate in which case great deviations occur the predicted results.

In this work, a data-driven convolution based neural network named inverse network is proposed to restore initial pulse through its nonlinear evolution. To confirm the wide adaptability of the proposed network, two independent typical nonlinear dynamics, i.e., the pulse evolution in fiber optical parametric amplifier (FOPA) systems and the soliton pair evolution in high-nonlinear fiber (HNLF) are discussed in detail [9]. We investigate the inverse network in various cases to ensure the precise predicting of initial pulse with different key parameters. Additionally, we evaluate the inverse network in case of different input sizes and different signal-to-noise ratios (SNRs). The results show that our method reveals a significant approach to estimating and optimizing the initial pulse in the FOPA and soliton pair systems. The paper is organized as follows. Section 2 illustrates the theory basis and the structure of proposed network. In section 3, the prediction results under two cases are given and the stability of the inverse network is discussed. Conclusions come in Section 4.

## 2. Theory and model

### 2.1. Pulse evolution in FOPA system

Single-pump FOPA system based on degenerated four-wave mixing of a pump wave, a signal wave and a conjugate idle wave in optical fibers can be described by coupled NLSEs [29]

$$\frac{\partial A_p}{\partial z} = -\frac{i}{2}\beta_{2p}\frac{\partial^2 A_p}{\partial^2 T} - \frac{1}{2}\alpha_p A_p + i\gamma[(|A_p|^2 + 2(|A_s|^2 + |A_I|^2))A_p + 2A_s A_I A_p^* \exp(i\Delta\beta z)]. \quad (1)$$

$$\frac{\partial A_s}{\partial z} = -d_s \frac{\partial A_s}{\partial T} - \frac{i}{2}\beta_{2s}\frac{\partial^2 A_s}{\partial^2 T} - \frac{1}{2}\alpha_s A_s + i\gamma[(|A_s|^2 + 2(|A_I|^2 + |A_p|^2))A_s + A_I^* A_p^2 \exp(-i\Delta\beta z)]. \quad (2)$$

$$\frac{\partial A_I}{\partial z} = -d_I \frac{\partial A_I}{\partial T} - \frac{i}{2}\beta_{2I}\frac{\partial^2 A_I}{\partial^2 T} - \frac{1}{2}\alpha_I A_I + i\gamma[(|A_I|^2 + 2(|A_s|^2 + |A_p|^2))A_I + A_s^* A_p^2 \exp(-i\Delta\beta z)], \quad (3)$$

where,  $A_p$ ,  $A_s$  and  $A_I$  are the slowly varying complex amplitude of pump wave, signal wave and idler wave, respectively.  $\alpha_i = \alpha$ ,  $i = p, s, I$  is the fiber loss coefficient.  $d_s$ ,  $d_I$  are the walk-off parameter for signal waves and idler waves.  $\Delta\beta = \beta_s + \beta_I - 2\beta_p$  is the linear phase mismatch

factor.  $T$  is measured in a frame of reference moving with the pulse at the group velocity and walk-off parameter.  $\gamma$  is nonlinear coefficient.

We consider the typical parameters of a FOPA system [30–31],  $d_s$  and  $d_l$  are set to -30 ps/km,  $\beta_{2s}$  equals to  $-\beta_{2l}$ , which is -20 ps<sup>2</sup>/km,  $\beta_{2p}$  is 0 ps<sup>2</sup>/km,  $\gamma$  is set to 17 km<sup>-1</sup>W<sup>-1</sup>,  $\alpha$  is set to 0.05 dB/km. Other parameters necessary in the simulation are listed as follow: signal wavelength is 1556 nm, pump wavelength is 1549 nm, second derivative of propagation constant is -1.2 ps<sup>2</sup>/km and fourth derivative of propagation constant is  $-2 \times 10^{-4}$  ps<sup>4</sup>/km. It should be mentioned that the zero-dispersion wavelength (ZDW) of the fiber is 1549 nm, and we set the pump wavelength at the ZDW, i.e., the  $\beta_{2p}$  is zero. In this work, amplitude of pulsed signal waves is given by

$$A_s(t, 0) = \sqrt{P_s} \exp(a_0(\frac{t^2}{T_0^2})), \quad (4)$$

where,  $P_s$  is the peak power of signal wave,  $T_0$  is the initial half width of signal pulse.  $a_0$  is coefficient of the Gaussian pulse with the value of -1.44.

## 2.2. Soliton pair evolution in HNLF

The evolution of soliton pair is considered as another typical nonlinear dynamic. In this work, the propagations of soliton pair with different initial relative amplitudes, separations, and phases are focused on. In this case, considering propagation attenuation, high-order group velocity dispersion, and fiber nonlinearity into account, the collision of two solitons can be governed by NLSE, which can be expressed as [9]

$$\frac{\partial A}{\partial z} + \frac{\alpha}{2}A + \beta_1 \frac{\partial A}{\partial t} - \sum_{k \geq 2} \frac{i^{k+1} \beta_k}{k!} \frac{\partial^k A}{\partial t^k} = i\gamma |A|^2 A, \quad (5)$$

where,  $A$  is the slowly varying complex amplitude of the soliton pair,  $\alpha$  is the fiber loss coefficient,  $\beta_k$  is group velocity dispersion of each order,  $\gamma$  is the nonlinear coefficient of fiber. Usually, Eq. (5) can be simplified as

$$\frac{\partial A}{\partial z} + \frac{\alpha}{2}A - \beta_2 \frac{\partial^2 A}{\partial t^2} = i\gamma |A|^2 A. \quad (6)$$

Corresponding parameters are set as:  $\beta_2$  is -20 ps<sup>2</sup>/km,  $\gamma$  is 17 km<sup>-1</sup>W<sup>-1</sup>,  $\alpha$  is 0.05 dB/km in the simulation. Soliton pair with different amplitudes and phases as the input of the nonlinear fiber is

$$A(0, t) = \sqrt{P} \left[ \text{sech}(\frac{t - q_0}{t_0}) + r \cdot \text{sech}(\frac{t + q_0}{t_0}) e^{i\theta} \right], \quad (7)$$

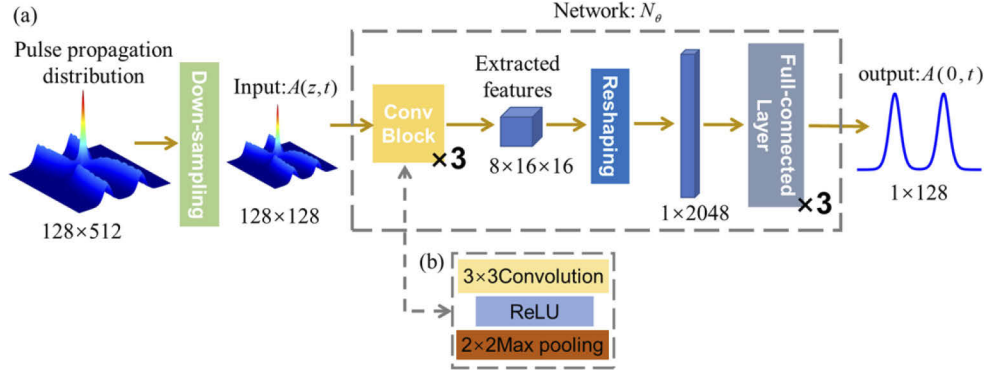
where,  $q_0$ ,  $r$  and  $\theta$  are the initial half separation, the relative amplitude, and the initial phase difference of two solitons, which greatly determines the characteristics of initial pulse and the evolution.  $t_0$  is the initial half width of pulse, which is set as 3 ps.  $P$  is the initial power of soliton pair and is set to 1 W.

The split-step Fourier (SSF) method is applied to solve the Eq. (1)–(3) and (6) iteratively, which is acknowledged as an effective and commonly used numerical approach. The detailed basis of SSF is discussed in the authority work [9].

## 2.3. Proposed data-driven inverse network structure

The schematic of the proposed data-driven inverse network architecture is shown in Fig. 1, which contains four parts: down-sampling, convolution blocks, reshaping, and full-connected layers. Now, let us take a closer look at the technical details of the inverse network. To lightweight the network and reduce computation demands, down sampling is firstly carried out. The primary

pulse propagation distribution is transformed to the input of inverse network, which is a matrix of discrete power profiles at different propagation distance  $A(z, t)$  of size  $m$  and  $n$  corresponding to the number of sampling points in time and distance domain. In this work,  $m$  and  $n$  are 128, which indicate that the size of input  $A(z, t)$  is  $128 \times 128$ .



**Fig. 1.** (a) Schematic of data-driven inverse network architecture: down-sampling, convolution blocks, reshaping and full-connected layers. (b) Structure of convolution block:  $3 \times 3$  convolution, rectified linear unit (ReLU) activation function and  $2 \times 2$  max pooling.

Here, convolution blocks are applied to extract the features of the input intensity matrix. The number of the convolution blocks is also flexible according to the size of the input image. In this work, three convolution blocks are used, which contain  $3 \times 3$  convolution, ReLU activation function and  $2 \times 2$  max pooling. Afterwards, the extracted feature cube of  $8 \times 16 \times 16$  size is reshaped to one dimensional feature vector ( $1 \times 2048$ ) and served as the input of three full-connected layers whose function are to further predict the initial state  $A(0, t)$ . Note that the reshaping process can be realized by other structures such as spatial pyramid pooling. We adopt a concise approach for the sake of computational efficiency while maintaining the accuracy of the prediction. The mapping function of the inverse network  $N_\theta$  can be solved by

$$N_\theta^* = \arg \min_{\theta \in \Theta} \|N_\theta(A(z, t)) - A_0(0, t)\|, \quad (8)$$

$$\forall (A(z, t), A_0(0, t)) \in D_T,$$

where,  $N_\theta$  is the mapping function of the neural network defined by a set of weights and biases  $\theta \in \Theta$ .  $N_\theta^*$  is the optimal solution of  $N_\theta$ .  $D_T$  is the training set.

$$D_T = \{(A(z, t)_i, A_0(0, t)_i); i = 1, \dots, K\}, \quad (9)$$

where,  $A(z, t)_i$  and  $A_0(0, t)_i$  are inputted intensity matrix and ground truth of initial pulse.  $K$  is the size of train set.

The weights and biases of proposed inverse network are updated by the backpropagation of root mean squared error (RMSE) between the predicted intensity of initial pulse and the ground truth and Adam is employed as the strategy. The program is implemented with Pytorch framework.

### 3. Results and discussion

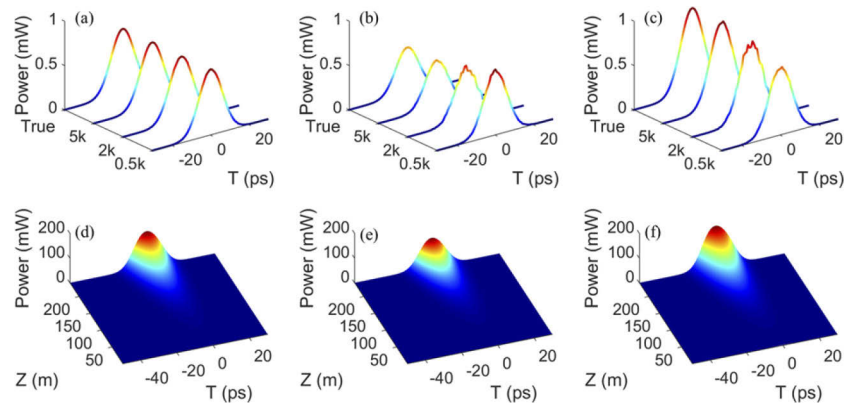
#### 3.1. Pulse prediction in the evolution of pulses in FOPA system

Before we demonstrated the prediction performance of proposed inverse network based on the mentioned data sets, details of training and testing process are discussed firstly to give an insight into how the network operated. The signal pulse prediction in FOPA is considered as two cases

of varying the initial signal power  $P_s$  and the initial half width of signal pulse  $T_0$ . In the first case,  $P_s$  is ranging from 0.5 mW to 1.0 mW with  $T_0$  remaining 20 ps. And  $T_0$  varies from 30 ps to 10 ps with  $P_s$  remaining 0.64 mW in the latter attempt. All the datasets are generated by SSF method in Python 3.7 platform. The computation platform owns an Intel 9900k CPU and a 2080Ti GPU.

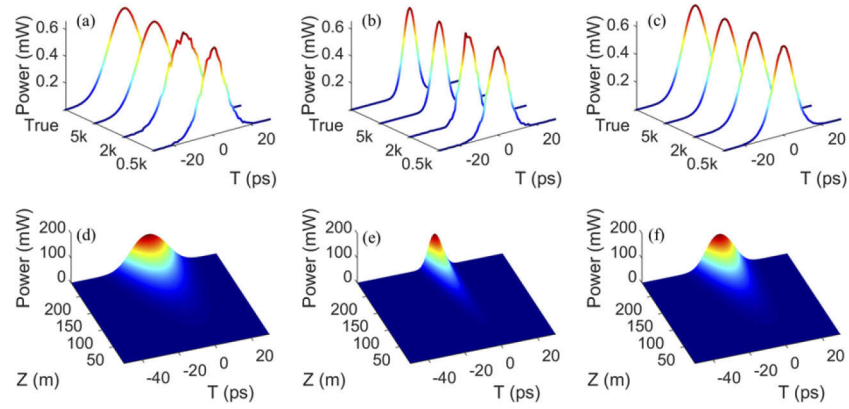
In each case, 1000 NLSE realizations are achieved by simulating the signal pulse propagations at the fiber length from 100 m to 228 m in the FOPA system. Thereinto, 900 realizations are used for training the network and 100 realizations are utilized in the testing phase. The cross-validation is used to overcome overfitting in the training phase. The proposed network is trained and tested ten times in each case. And each time the training set and the testing set are randomly split, i.e., the testing set is independent of the training set in the whole training phase. The final prediction accuracy is the average of ten outcomes. It should be mentioned that the proposed inverse network can obtain a similar performance on astringency and prediction accuracy on the dataset with different propagations distances. Hence, without loss of generality, the pulse propagation distributions from 100 m to 228 m are utilized in the training and testing phases to show a general prediction result.

To begin with, the case of different initial signal power is considered. The predicted power profiles and the ground truth setting with initial power of 0.735 mW, 0.526 mW and 0.969 mW with corresponding epochs of 500, 2000, and 5000 are illustrated in Fig. 2(a)-(c) as typical examples. Corresponding pulse propagations are shown in Fig. 2(d)-(f). As shown in Fig. 2(d)-(f), initial signal power greatly affects the temporal amplitude of pulse propagation, especially in the late stage of evolution. It is clearly seen in Fig. 2(a)-(c), the prediction results are closer to the ground truth, as the training goes on. During the epoch number increases to 2000, the shapes of predicted initial pulse gradually approximate the target one. Admirable consistencies are reached after 5000 epochs training, whether setting the initial signal power low or high, which indicates the great performance on the different initial power case.



**Fig. 2.** (a), (b) and (c) Predicted power distribution profiles with initial signal power of 0.735 mW, 0.526 mW and 0.969 mW at corresponding epochs of 500, 2000, 5000 and ground truth. (d), (e) and (f) Corresponding pulse propagation process of (a), (b) and (c).

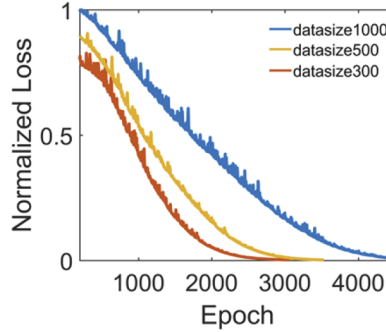
Additionally, we discuss the case with various initial pulse widths. Three representative predicted power profiles are illustrated in Fig. 3(a)-(c). Corresponding pulse propagations are shown in Fig. 3(d)-(f). Different from Fig. 2(d)-(f), the initial pulse width mainly influences the broadening of the pulse and has slight impact on the temporal amplitude during the propagation, which is illustrated in Fig. 3(d)-(f). The predicted pulses with different epochs show a similar quality to Fig. 2(a)-(c). The inverse network can output strongly consistent initial pulse distribution when the epoch grows to 5000.



**Fig. 3.** (a), (b) and (c) Predicted power distribution profiles with initial signal half pulse width of 27.10 ps, 10.88 ps and 20.64 ps at corresponding epochs of 500, 2000, 5000 and ground truth. (d), (e) and (f) Corresponding pulse propagation process of (a), (b) and (c).

The number of epochs that obtain a good prediction result can be floating, which is related to learning rate, size of the dataset and other network parameters. In this work, thanks to the mentioned network lightweight approaches, a 5000-epoch training only costs approximately 30 seconds.

We further investigate the relation between dataset size and required training epoch. The dataset size is changed from 1000 to 500 and 300, and the network setting remains the same. The network is trained and tested under different dataset sizes according to the case in Fig. 3. The normalized losses of testing set under different dataset sizes are plotted in Fig. 4. The results show that reducing the dataset size can speed up the convergence of network. However, reducing the dataset size will decline the effective prediction range of initial pulse.



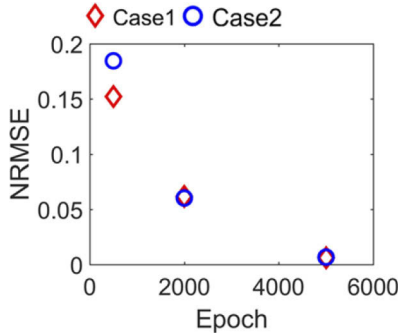
**Fig. 4.** Normalized loss of testing set with different dataset sizes.

Here, the normalized root mean squared error  $R_1$  is applied to evaluate the performance of the inverse network, which is as follow

$$R_1(x, \hat{x}) = \sqrt{\frac{\sum (x_0 - \hat{x}_0)^2}{\sum (x_0)^2}}, \quad (10)$$

where  $\hat{x}_0$  and  $x_0$  are the predicted initial power distribution by the inverse network and ground truth. Therefore, we measured  $R_1$  at different training epoch in two cases (i.e., Case 1: varying initial signal power, Case 2: varying initial half width of signal wave) on the 100 different pulse

propagation maps, which is displaced in Fig. 5. The  $R_1$  on testing sets further establish the great predicting performance of the proposed method. After 5000 epochs training, the  $R_1$  on two testing sets are 0.0063 and 0.0067, respectively. The high-degree of consistency on both variable power case and variable width case demonstrate that proposed inverse network can predict the initial input state with a limited pulse propagation in FOPAs.



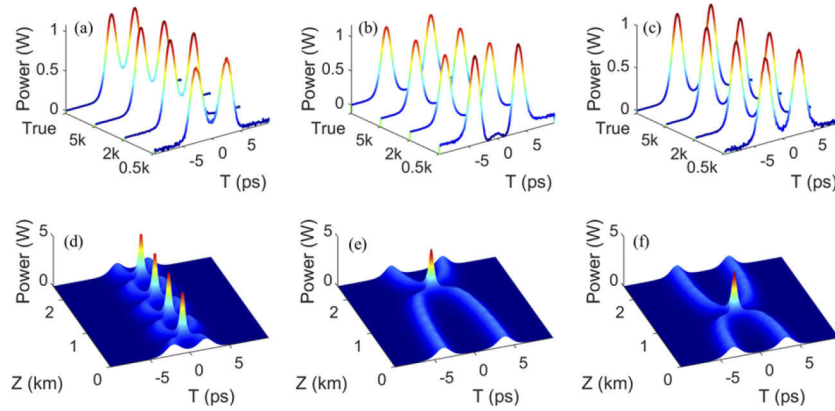
**Fig. 5.** Normalized root mean squared errors on two cases with corresponding epochs of 500, 2000, 5000.

### 3.2. Pulse prediction in the soliton pair evolution in HNLF

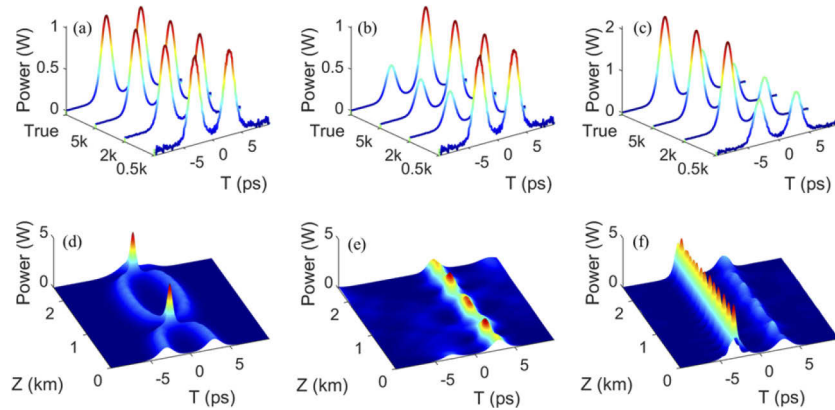
In addition, we consider pulse prediction in a more complex dynamic that the soliton pair evolution in HNLF. Two different cases are analyzed that varying the initial half separation  $q_0$  and the relative amplitude  $r$  of two solitons. In two conditions,  $q_0$  is set from 2 ps to 4 ps with  $r$  of 1, and  $r$  is set as 0.5 to 1.5 while  $q_0$  remaining 3 ps, respectively. Similarly, in each case, 1000 NLSE realizations are obtained by simulating the nonlinear propagations of soliton pair at distance of 500 m to 1140 m in HNLF. Thereinto, 900 realizations are used for training the network and 100 realizations are applied in the test phase.

To begin with, we discuss the three typical prediction results of different  $q_0$  with training epochs of 500, 2000 and 5000 in the condition of relative amplitude and relative phase setting to 1 and 0 respectively, which are plotted in Fig. 6(a)-(c). The corresponding pulse propagation process is shown in Fig. 6(d)-(f). The dynamic of soliton pair with the same power and phase is a periodic attraction that the soliton pair attracts to each other, and periodic collision occurs along with the propagation distance. The  $q_0$  affects the periodic length that the larger  $q_0$ , the longer a collision occurs. At the epoch of 500, it is observed in Fig. 6(a)-(c) that the generated initial pulses perform poorly and distinct with the ground truth. As the number of epochs turns to 2000, the distributions of prediction pulse are approximate to the target ones. The inverse network can generate a highly exact estimation of initial waveform when the epoch number increases to 5000.

Then, we test the prediction performance in the case of different  $r$  while the initial half separation remaining 3 ps. The initial relative phase of two solitons is 0. As established in Fig. 7(a)-(c), we select three typical predictions of high, low, and normal relative amplitude  $r$  when training epochs reach 500, 2000 and 5000. Corresponding pulse propagation process of Fig. 7(a)-(c) are given in Fig. 7(d)-(f), in which the soliton pair attract to each other and periodic crash when the relative amplitude is near 1. Once there is a great amplitude difference between soliton pair, two solitons propagate separately along with the fiber length and nonlinear fluctuations occur on both solitons. Besides, the proposed inverse network reveals a similar performance during training compared to the above case of different  $q_0$ . After 2000 epochs of training, the generated initial waveform gradually matches the expected distribution from a badly agreed primary prediction. Exact initial pulse distributions are obtained with their sectional evolution maps when the number of epochs turns to 5000, as clearly shown in Fig. 7(a)-(c).



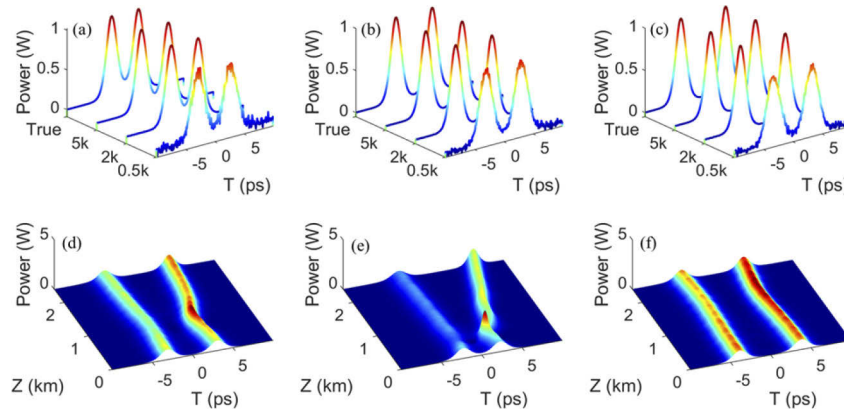
**Fig. 6.** (a), (b) and (c) Predicted power distribution profiles with initial half separation of 2.078 ps, 3.862 ps and 2.922 ps at corresponding epochs of 500, 2000, 5000 and ground truth. (d), (e) and (f) Corresponding pulse propagation process of (a), (b) and (c).



**Fig. 7.** (a), (b) and (c) Predicted power distribution profiles with relative amplitude of 1.005, 0.640 and 1.422 at corresponding epochs of 500, 2000, 5000 and ground truth. (d), (e) and (f) Corresponding pulse propagation process of (a), (b) and (c).

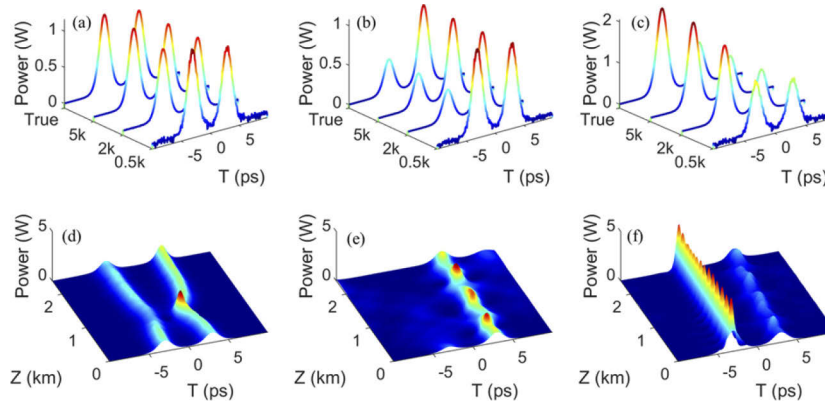
Note that the initial phase difference of neighboring solitons  $\theta$  is also a non-negligible factor that impacts the evolution of soliton pair. The in-phase cases are discussed above. We further test the proposed network under two aforesaid cases by interchanging the initial phase difference  $\theta$  to  $\pi/4$  for verifying the robustness of the proposed method. The case of variable  $q_0$  is firstly given in Fig. 8(a)-(c). Comparing Fig. 6(d)-(f) to Fig. 8(d)-(f), the difference of initial phase of two solitons  $\theta$  leading to wide divergences of soliton pair propagation. Different from the periodic attraction in Fig. 6(d)-(f), soliton pair appear diverse propagation process when the relative phase of  $\pi/4$ . In Fig. 8(d) and (e), the two solitons repulse strongly and the distance between them increases during subsequent evolution, after an initial phase of attraction. However, weakening interaction between the two solitons to make solitons propagating smoothly in a long distance is essential in the optical communication system. The expected propagation process can be achieved by setting the appropriate  $q_0$ , as demonstrated in Fig. 8(f).

Furthermore, we discuss the predictions with  $\pi/4$  initial phase difference and various relative amplitude. Three traditional facts with  $r$  of 1.036, 0.658 and 1.432 are shown in Fig. 9(a)-(c). Corresponding evolutions of soliton pair are shown in Fig. 9(d)-(f). Similar to the results of the



**Fig. 8.** (a), (b) and (c) Predicted power distribution profiles with relative phase and initial half distance of 2.078 ps, 3.862 ps and 2.922 ps at corresponding epochs of 500, 2000, 5000 and ground truth. (d), (e) and (f) Corresponding pulse propagation process of (a), (b) and (c).

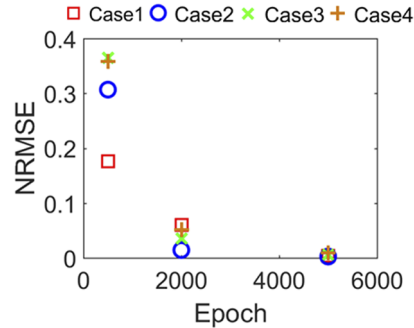
previous cases, accurate initial pulse waveforms are generated by the inverse network after 5000 epochs training. Comparing Fig. 7(d) to Fig. 9(d), the relative phase has a great impact on the behavior of soliton pair when relative amplitude is about 1. When the initial phase difference of two solitons is set to  $\pi/4$ , instead of periodically attracting, solitons firstly go through the attraction phase then attempt to repel and separate. However, if the amplitude of two solitons is largely different, the influence of relative phase on the soliton pair propagation is tiny, the behaviors of soliton pair are semblable with Fig. 7(e)-(f).



**Fig. 9.** (a), (b) and (c) Predicted power distribution profiles with  $\pi/4$  relative phase and relative amplitude of 1.036, 0.658 and 1.432 at corresponding epochs of 500, 2000, 5000 and ground truth. (d), (e) and (f) Corresponding pulse propagation process of (a), (b) and (c).

Afterwards, the normalized root mean squared error  $R_1$  is utilized in the four concerned cases, which are listed as Case 1: varying initial half separation with initial phase difference of 0, Case 2: varying relative amplitude with initial phase difference of 0, Case 3: varying initial half separation with initial phase difference of  $\pi/4$  and Case 4: varying relative amplitude with initial phase difference of  $\pi/4$ , respectively. The calculated  $R_1$  are demonstrated in Fig. 10. As the number of epochs grows to 5000, the  $R_1$  with four cases are 0.0037, 0.0025, 0.0062, and 0.0096

respectively, which verify that the predicted initial distributions are highly consistent with the ground truth in entire testing sets.



**Fig. 10.** Normalized root mean squared errors on four cases with corresponding epochs of 500, 2000, 5000.

### 3.3. Stability of the inverse network

Furthermore, we investigate the stability of the inverse network. The performance of the inverse network to the noise is discussed firstly. Gaussian noises are added to the testing sets in previous cases (i.e., two FOPA cases and four soliton cases). The inverse network is examined on the corresponding testing sets with the SNR of 10 dB, 20 dB, and 30 dB, respectively. The corresponding normalized root mean squared errors are listed in Table 1 and 2, respectively.

**Table 1. Normalized root mean squared errors in FOPA cases with different SNRs**

SNR	Case 1	Case 2
10 dB	0.0749	0.0749
20 dB	0.0278	0.0225
30 dB	0.0085	0.0098

The results show that the inverse network is robust to the noise. When the SNR of the testing set is 30 dB, the predicting accuracy is hardly influenced. As the SNR decreases, the deviation of predicted pulse to the truth increases.

Table 2 further demonstrates the noise immunity of the inverse network. When the SNR of the testing sets is higher than 20 dB, the  $R_1$  is still very low, which indicates that stability of the inverse network to Gaussian noise is significant. However, in case that the SNR of testing set is about 10 dB, the prediction error has increased substantially. The network needs to be retrained to accommodate the noise data. Besides, the  $R_1$  in the soliton cases are lower than that of FOPA cases at same SNR. The propagation characteristics that the amplitude of soliton is relatively stable compared with that of FOPA system can lead to the advantage. In FOPA system, the amplitude of signal pulse is weak in the beginning, high noise influences the early stage of the FOPA evolution greatly, thus leading to higher deviation on the prediction initial pulse.

Despite the stability of network to noise, the performance of inverse network on different sizes of input is also studied. We emphasize the prediction by using less or even a dozen discrete power distributions profiles at different propagation locations.

In detail, the datasets with different sizes are obtained by discrete sampling step size over propagation distance. In two discussed nonlinear dynamics, the origin step size is set as 1 m and 5 m. We ulteriorly considered step size of 2 m, 4 m, 8 m within FOPA cases and step size

**Table 2. Normalized root mean squared errors in soliton cases with different SNRs**

SNR	Case 1	Case 2	Case 3	Case 4
10 dB	0.0428	0.0269	0.0515	0.0801
20 dB	0.0125	0.0096	0.0296	0.0223
30 dB	0.0047	0.0027	0.0161	0.0104

of 10 m, 20 m, 40 m in the soliton cases, which corresponds to 64, 32, 16 power profiles. We keep the structure of the inverse network as well as the realization of network training and testing unchanged. The results of two FOPA cases are shown in Table 3.

**Table 3. Normalized root mean squared errors in FOPA cases**

Number of profiles (Epoch)	Case 1	Case 2
64 (5k)	0.0087	0.0092
32 (5k)	0.0118	0.0125
16 (5k)	0.0179	0.0152

Similar prediction performance on the whole testing set is obtained compared to results shown in Fig. 4. Despite the decreasing of power profiles used in training having a minor influence on the  $R_1$ , the inverse network reaches a great prediction accuracy when the epoch grows to 5000. Especially when the power distribution profiles are reduced to an eighth of the original setting, the  $R_1$  is still below 0.02. Besides, the  $R_1$  on the soliton pair evolution cases are displaced in Table 4.

**Table 4. Normalized root mean squared errors in soliton cases**

Number of profiles (Epoch)	Case 1	Case 2	Case 3	Case 4
64 (5k)	0.0071	0.0108	0.0087	0.0095
32 (5k)	0.0148	0.0097	0.0125	0.0134
16 (5k)	0.0225	0.0178	0.0189	0.0154

Compared with the prediction based on 128 distance discrete power distributions that given in Fig. 9, the  $R_1$  slightly increases when the number of power distribution profiles is declined at epoch of 5000 in Table 4. However, the increase is acceptable since for the  $R_1$  is lower than 0.025 after 5000 epochs training when the input size is heavily reduced. The results show that decreasing the number of power distribution profile used in training has a minor effect on the convergence rate and prediction accuracy in both nonlinear dynamics, which further reveals the robustness of the proposed network in a small input size.

In summary, the comparisons of normalized root mean squared errors on all tests performed are detailed given in Table 5. The discussion of adaptability on different noise levels and input sizes proves the inverse network a promising approach to assess the initial state of nonlinear fiber system. In addition, the proposed network can be very concise and timesaving by virtue of a series of lightweight methods, with potential application in the real-time system optimization and pulse prediction.

**Table 5. Comparison of normalized root mean squared errors on all tests performed**

	Different SNRs with input size of 128			Different input sizes without noise		
	30 dB	20 dB	10 dB	64	32	16
FOPA Case1	0.0085	0.0278	0.0749	0.0087	0.0118	0.0179
FOPA Case2	0.0098	0.0225	0.0749	0.0092	0.0125	0.0152
Soliton Case1	0.0047	0.0125	0.0428	0.0071	0.0148	0.0225
Soliton Case2	0.0027	0.0096	0.0269	0.0108	0.0097	0.0178
Soliton Case3	0.0161	0.0296	0.0515	0.0087	0.0125	0.0189
Soliton Case4	0.0104	0.0223	0.0801	0.0095	0.0134	0.0154

#### 4. Conclusions

The initial pulse state has a great impact on the nonlinear pulse evolution in fiber optics, which makes it significant to assess the initial distribution of a nonlinear optical fiber system. In this work, an inverse network that contains down-sampling, convolution blocks, reshaping, and full-connected layers is proposed to predict the initial pulse distribution of two typical nonlinear dynamics (i.e., the pulse evolution in FOPA system and the soliton pair evolution in HLNF) through their limited power profiles at different propagation distance. When the epoch grows to 5000, the inverse network can output accurate predicted initial pulse waveform in both cases. While maintaining the prediction accuracy, the highly compressed inverse network demonstrates great performance in efficiency. We further test the inverse network with datasets of different SNRs and input sizes. The inverse network shows pretty stability to the deviation on the testing set. The proposed inverse network opens new perspectives for initial pulse estimation and pulse parameter optimization for future nonlinear fiber optics systems.

**Funding.** National Key Research and Development Program of China (2019YFB1803500); Fundamental Research Funds for the Central Universities (2682021GF018); Foundation of Science and Technology Department of Sichuan Province, China (2020YJ0016).

**Disclosures.** The authors declare no conflicts of interest.

**Data Availability.** This data, models, or code generated or used are available from the corresponding author by reasonable request.

#### References

1. H. M. Masoudi, "A novel nonparaxial time-domain beam-propagation method for modeling ultrashort pulses in optical structures," *J. Lightwave Technol.* **25**(10), 3175–3184 (2007).
2. F. Poletti and P. Horak, "Description of ultrashort pulse propagation in multimode optical fibers," *J. Opt. Soc. Am. B* **25**(10), 1645–1654 (2008).
3. H. Steffensen, J. R. Ott, K. Rottwitz, and C. J. McKinstrie, "Full and semi-analytic analyses of two-pump parametric amplification with pump depletion," *Opt. Express* **19**(7), 6648–6656 (2011).
4. I. Cristiani, L. Tartara, G. P. Banfi, and V. Degiorgio, "Ultrashort-pulse investigation of the propagation properties of the LP11 mode in 1.55- $\mu$ m communication fibers," *Opt. Lett.* **26**(22), 1758–1760 (2001).
5. D. Bigourd, L. Lago, A. Kudlinski, E. Hugonnot, and A. Mussot, "Dynamics of fiber optical parametric chirped pulse amplifiers," *J. Opt. Soc. Am. B* **28**(11), 2848–2854 (2011).
6. L. F. Zhang, C. X. Li, H. Z. Zhong, C. W. Xu, D. J. Lei, Y. Li, and D. Y. Fan, "Propagation dynamics of super-Gaussian beams in fractional Schrödinger equation: from linear to nonlinear regimes," *Opt. Express* **24**(13), 14406–14418 (2016).
7. M. Antonelli, A. Shtaiif, and Mecozzi, "Modeling of nonlinear propagation in space-division multiplexed fiber-optic transmission," *J. Lightwave Technol.* **34**(1), 36–54 (2016).
8. X. Q. Zhong, L. F. Chen, K. Cheng, N. Yao, and J. N. Sheng, "Generation of single or double parallel breathing soliton pairs, bound breathing solitons, moving breathing solitons, and diverse composite breathing solitons in optical fibers," *Opt. Express* **26**(12), 15683–15692 (2018).
9. G. P. Agrawal, *Nonlinear Fiber Optics* (Academic, 2013), Chap. 5-7.
10. L. Salmela, N. Tsipinakis, A. Foi, C. Billet, J. M. Dudley, and G. Genty, "Predicting ultrafast nonlinear dynamics in fibre optics with a recurrent neural network," *Nat Mach Intell* **3**(4), 344–354 (2021).

11. G. Genty, L. Salmela, J.M. Dudley, D. Brunner, A. Kokhanovskiy, S. Kobtsev, and S. K. Turitsyn, "Machine learning and applications in ultrafast photonics," *Nat. Photonics* **15**(2), 91–101 (2021).
12. F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks," *IEEE Commun. Surv. Tutorials* **21**(2), 1383–1408 (2019).
13. S. H. Wang, S. Y. Xiang, G. Q. Han, Z. W. Song, Z. X. Ren, A. J. Wen, and Y. Hao, "Photonic associative learning neural network based on VCSELs and STDP," *J. Lightwave Technol.* **38**(17), 4691–4698 (2020).
14. Y. N. Han, S. Y. Xiang, Z. X. Ren, C. T. Fu, A. J. Wen, and Y. Hao, "Delay-weight plasticity-based supervised learning in optical spiking neural networks," *Photon. Res.* **9**(4), B119–127 (2021).
15. Y. F. Zhang, L. Yu, Z. L. Hu, L. Cheng, H. Sui, H. N. Zhu, G. M. Li, B. Luo, X. H. Zou, and L. S. Yan, "Ultrafast and accurate temperature extraction via kernel extreme learning machine for BOTDA sensors," *J. Lightwave Technol.* **39**(5), 1537–1543 (2021).
16. M. Närhi, L. Salmela, J. Toivonen, C. Billet, J. M. Dudley, and G. Genty, "Machine learning analysis of extreme events in optical fibre modulation instability," *Nat. Commun.* **9**, 4923 (2018).
17. T. Baumeister, S. L. Brunton, and J. N. Kutz, "Deep learning and model predictive control for self-tuning mode-locked lasers," *J. Opt. Soc. Am. B* **35**(3), 617–626 (2018).
18. U. Teğin, N. Ul. Dinç, C. Moser, and D. Psaltis, "Reusability report: predicting spatiotemporal nonlinear dynamics in multimode fibre optics with a recurrent neural network," *Nat Mach Intell* **3**(4), 344–354 (2021).
19. D. S. Wang, Y. C. Song, J. Li, J. Qin, T. Yang, M. Zhang, X. Chen, and A.C. Boucouvalas, "Data-driven optical fiber channel modeling: a deep learning approach," *J. Lightwave Technol.* **38**(17), 4730–4743 (2020).
20. X. Jiang, D. Wang, Q. Fan, M. Zhang, C. Lu, and A. P. Tao Lau, "Solving the nonlinear Schrödinger equation in optical fibers using physics-informed neural network," *2021 Optical Fiber Communications Conference and Exhibition (OFC)*, 1–3 (2021).
21. R. Q. Wang, L. M. Ling, D. L. Zeng, and B.F. Feng, "A deep learning improved numerical method for the simulation of rogue waves of nonlinear Schrödinger equation," *Communications in Nonlinear Science and Numerical Simulation*, **101**, 105896 (2021).
22. S. Boscolo, J. M. Dudley, and C. Finot, "Modelling self-similar parabolic pulses in optical fibres with a neural network," *Results in Optics* **3**, 100066 (2021).
23. S. Boscolo and C. Finot, "Artificial neural networks for nonlinear pulse shaping in optical fibers," *Opt. Laser Technol.* **131**, 106439 (2020).
24. N. Gautam, A. Choudhary, and B. Lall, "Comparative study of neural network architectures for modelling nonlinear optical pulse propagation," *Opt. Laser Technol.* **64**, 102540 (2021).
25. Q. R. Fan, G. Zhou, T. Gui, L. Chao, and A. P. T. Lau, "Advancing theoretical understanding and practical performance of signal processing for nonlinear optical communications through machine learning," *Nat. Commun.* **11**(1), 3694 (2020).
26. S. L. Zhang, F. Yaman, K. Nakamura, T. Inoue, V. Kamalov, L. Jovanovski, V. Vusirikala, E. Mateo, Y. Inada, and T. Wang, "Field and lab experimental demonstration of nonlinear impairment compensation using neural networks," *Nat. Commun.* **10**(1), 3033 (2019).
27. O. Sidelnikov, A. Redyuk, S. Sygletos, M. Fedoruk, and S. Turitsyn, "Advanced convolutional neural networks for nonlinearity mitigation in long-haul WDM transmission systems," *J. Lightwave Technol.* **39**(8), 2397–2406 (2021).
28. R. A. Herrera, "Evaluating a neural network and a convolutional neural network for predicting soliton properties in a quantum noise environment," *J. Opt. Soc. Am. B* **37**(10), 3094–3098 (2020).
29. S. Wabnitz and J. M. Soto-Crespo, "Stable coupled conjugate solitary waves in optical fibers," *Opt. Lett.* **23**(4), 265–267 (1998).
30. Y. Z. Li, L. J. Qian, D. Q. Lu, and D. Y. Fan, "Ultrafast four-wave mixing in single-pumped fibre optical parametric amplifiers," *J. Opt. A: Pure Appl. Opt* **8**(8), 689–694 (2006).
31. X. M. Liu, "Theory and experiments for multiple four-wave-mixing processes with multifrequency pumps in optical fibers," *Phys. Rev. A* **77**(4), 043818 (2008).